EOS ClearingHOuse (ECHO) Provider Training Class

Keith Wichmann
ECHO System Engineer/Architect
wichmann@gst.com
(240) 542-1117

Robin Pfister
ESDIS
Robin.Pfister@nasa.gov
(301) 614-5171

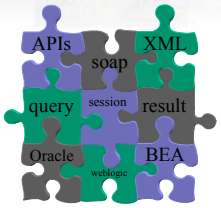
Global Science & Technology, Inc.

Agenda

Monday, November 3rd, 2003 – Provider Class				
11/3/2003	8:30 AM	0:30	9:00 AM	Coffee
11/3/2003	9:00 AM	0:30	9:50 AM	ECHO Overview
11/3/2003	9:50 AM	0:10	10:00 AM	Registration
11/3/2003	10:00 AM	0:15	10:15 AM	Roles in ECHO
11/3/2003	10:15 AM	0:15	10:30 AM	Provider Acct Service
11/3/2003	10:30 AM	0:15	10:45 AM	Break
11/3/2003	10:45 AM	1:00	11:45 AM	Ingest
11/3/2003	11:45 AM	1:30	1:15 PM	Lunch
11/3/2003	1:15 PM	1:20	2:35 PM	Orders / Options / POMS
11/3/2003	2:35 PM	0:10	2:45 PM	Provider Policies
11/3/2003	2:45 PM	0:15	3:00 PM	Break
11/3/2003	3:00 PM	1:30	4:30 PM	Protecting Data
11/3/2003	4:30 PM	0:30	5:00 PM	Ops Team Support

ECHO Overview

- What is ECHO?
- Goal of ECHO
- Goal of Workshop
- History of ECHO
- Guidance/Advisory Groups
- Architecture Overview
- Current State of ECHO



What is ECHO?

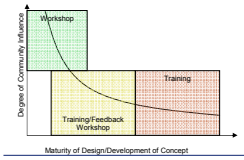
- Clearinghouse of spatio-temporal metadata.
- Order Broker
- User and provider account service
- Services clearinghouse and broker (future):
  - Data Services (e.g. Subsetting)
  - Search Services (e.g. Coincidence Searching)
  - Administrative Services (e.g. B&A)

Goal of ECHO

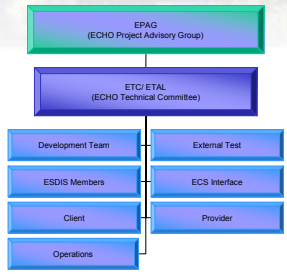
- Respond to user needs for more flexible interfaces for access to EOS data.
- Serve as a portal to Earth Science metadata.
- Allow providers of data to share their metadata and offload some of their search responsibilities.
- Broker orders from clients to the appropriate providers, providing tracking services for both the client and the provider.
- Present a messaging interface based on XML.
- NOT a GUI.

Goal of Training Workshop

- Train Provider Developers on the use of the ECHO Provider APIs.
- All-way, informal interaction among participants. We need your feedback.
- Familiarize Provider Developers with how to use the materials: web site and users manual.
- Collect any requirements we may have missed.



ECHO Committees



Guidance/Advisory Groups

- **ECHO Technical Committee:**
  - Originated to study IMS alternatives, independent of the ECS architecture.
  - Representatives from various contractors, ESDIS, and ECHO providers.
  - Continues to provide technical input into the direction of ECHO on a monthly basis.
- **Advisory Panel: made up of ESDIS and non-ESDIS representatives.**

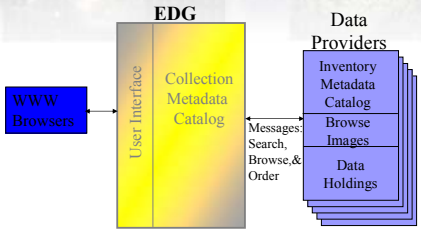
Architecture Overview:

System Drivers

- **Present an API for organizations to connect their own user interfaces and programs to**
- **Make it easy for providers of Earth Science data and services to participate in the system**
- **Provide searches that respond quickly**
- **Broker orders for both data and services**
- **Minimize operational costs**
- **Build upon advances in industry and use e-commerce systems as a model**
- **Build a system that can be scaled up to handle large numbers of requests**

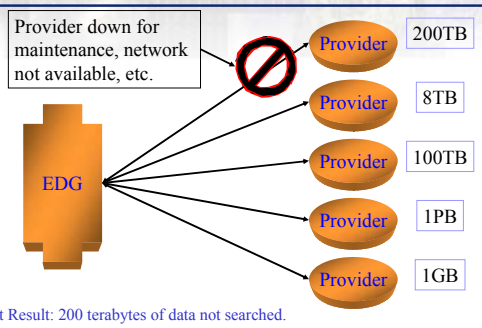
Architecture Overview:

Current Information Architecture



Architecture Overview:

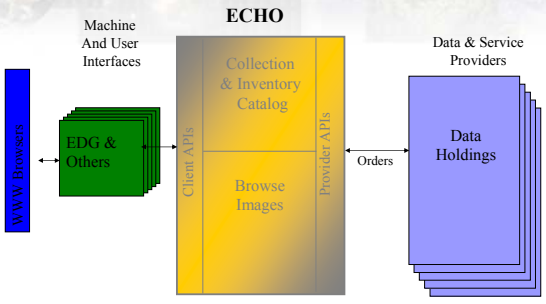
Current Metadata Access



Net Result: 200 terabytes of data not searched.
Search time dependent on lowest common denominator.

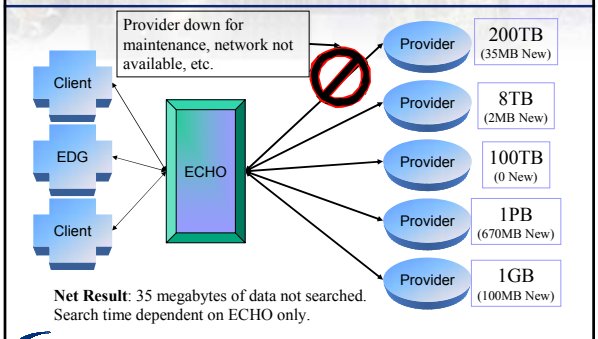
Architecture Overview:

Information Architecture w/ECHO



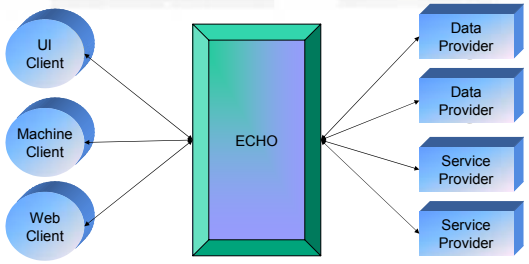
Architecture Overview:

ECHO Metadata Access

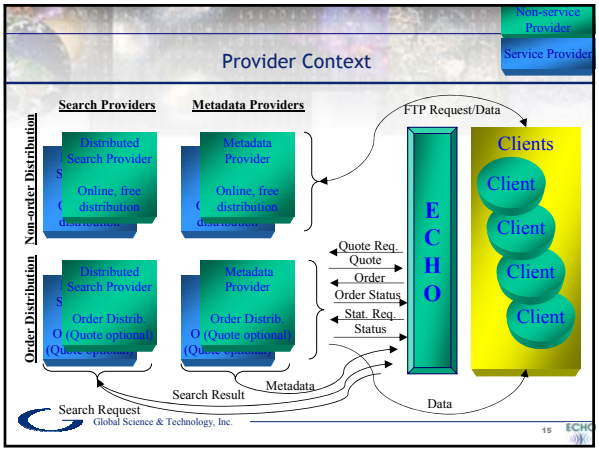


Net Result: 35 megabytes of data not searched.
Search time dependent on ECHO only.

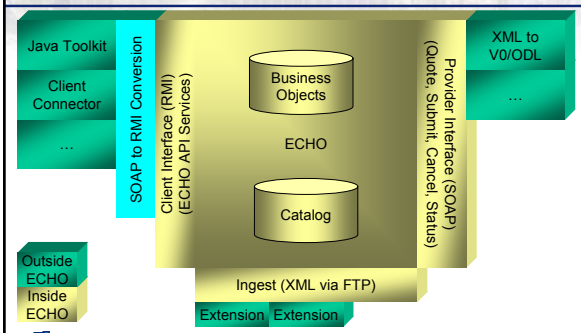
High Level ECHO Context



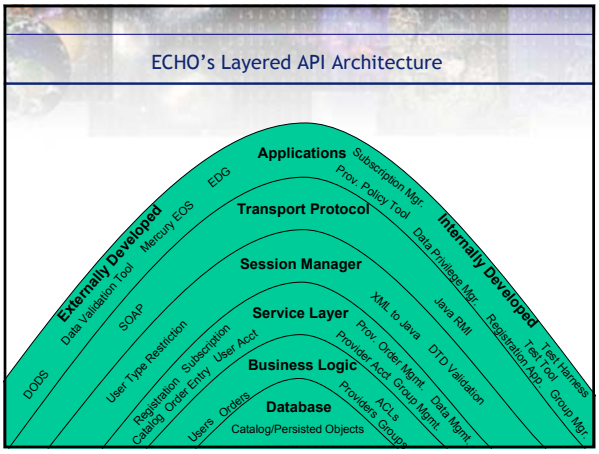
Provider Context



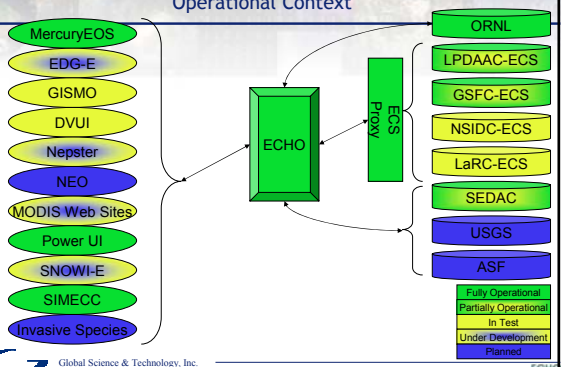
ECHO Interfaces



ECHO's Layered API Architecture



Operational Context



Current State of ECHO

- **Second Operational Version (Release 2 - Version 5.0):**
  - Functional and useful, but not complete or necessarily completely correct.
    - Vetted through previous workshops and a year of operations
    - Clients have provided feedback and will continue to
  - Changes are coming
    - Data Model Review changes will continue to cause some basic changes in the conceptual model of the system
    - XML has been leveraged to try to minimize impact of changes
    - Some structuring of the messages and tag names will be impacted
    - New functions will be added
    - Some parts of the model that are not being used will be removed
- **Your feedback is desired, welcome and needed!**
  - What will make your provider interface work better?
  - Are there missing requirements?

ECHO URLs

- **The Official ECHO Web Page**
  - http://www.echo.eos.nasa.gov
- **Ingests should be sent to**
  - Ingest.echo.eos.nasa.gov
- **The APIs are available at**
  - Api.echo.eos.nasa.gov
- **Browse imagery is available at**
  - Browse.echo.eos.nasa.gov
- **Unofficial web sites with useful information**
  - Canyon.gst.com/public
  - Dangermouse.gst.com/dmr (Data Model Review)

Vocabulary

- **ECHO – EOS Clearing HOuse**
- **EOS – Earth Observing System**
- **EJB – Enterprise Java Bean**
- **J2EE – Java 2 Enterprise Edition**
- **SOAP – Simple Object Access Protocol**
- **DTD – Data Type Definition**
- **XML – eXtensible Markup Language**

Understanding the ECHO APIs

Global Science & Technology, Inc.

Registration Service Transactions for Providers

The Registration Service allows a potential provider to apply to be an ECHO provider. ECHO operations staff will do some setup and configuration after the application is submitted.

Global Science & Technology, Inc.

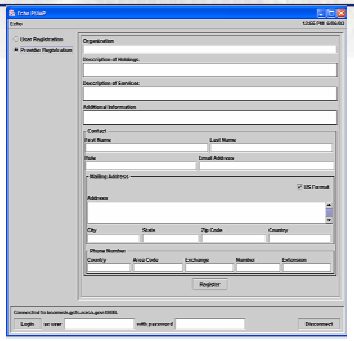
Provider Application Transaction

- **SubmitProviderApplicationRequest**
  - A potential provider places an application to become an ECHO provider through the Registration Service.
  - Application information includes organization name, provider contact information and some description of the provider's data holdings.
  - ECHO will respond with an XML message to indicate if the application is successful or not and assigns a temporary tracking id that the potential provider can use to communicate with the ECHO operations team.

Sample Provider Application Request

```
<RegistrationService>
        <ProviderApplicationRequest>
                <ProviderApplication>
                        <OrganizationName>Global Science and Technology</OrganizationName>
                        <ProviderContact>
                                <ContactRole>Manager</ContactRole>
                                <ContactFirstName>Paul</ContactFirstName>
                                <ContactLastName>Clemens</ContactLastName>
                        </ProviderContact>
                </ProviderApplication>
        </ProviderApplicationRequest>
</RegistrationService>
```


PUMP Provider Application Screen



Roles in ECHO

-or-

How to have a split personality and maintain your API…

(A brief description of the roles capability added into ECHO Version 5)

Global Science & Technology, Inc.

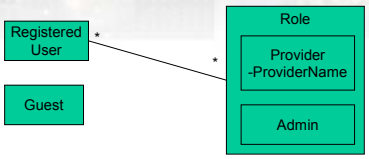
Two Types of Users

- **As of ECHO Version 5.0, there are two types of users**
  - Guest
    - This is a user of the system who does not call the login method of the Session Manager
    - This user is not guaranteed (or in most cases allowed) persistence of items in ECHO beyond their orders
  - Registered User
    - This is a user who has successfully called the login method of the Session Manager with their user id and password
    - This user can save and recall queries, result sets, persist account information, access a list of orders submitted or created but not submitted, etc.
- **Previously, there was a Provider User as well, but they may no longer log into the system**

Registered Users May Have Roles

- **Currently, there are two possible types of roles that may be given to a Registered User**
  - Note: A Registered User does not HAVE to have any roles
  - Administrator
    - A User with this role can access additional transactions that allow them to act as any user, delete users, and allow them to see any order. It does not allow them access to data that has been hidden via the DMS unless they act as a user that has access.
  - Provider
    - A User with this role can access additional transactions that allow them control over the Provider Account since no one can log in as the Provider anymore.
      - Provider Account Service
      - Provider Order Management Service
      - Data Management Service

Role



- **A Registered User can have many Roles, including Administrator, and Provider for any number of Registered Providers**
- **A Guest User cannot have any Roles**

Administrator Role

- **Once logged in, a User with this role acts as them self until they do a change context operation.**
- **If the user performs a change user context, then future transactions will operate as if the administrator is the user they changed context to**
  - The ECHO log files will show both the administrator's login name and the user they are acting as so that a record is maintained of who actually performed a function
- **If the administrator user performs a change provider context, then they will be allowed to access Provider only transactions, and will act as the specified Provider**
  - An admin user can change to any Provider

Provider Role

- **A User may have multiple Provider Roles, one for each Provider that they represent**
  - For instance, EDC has two separate data centers that they operate which may share personnel
- **Once a user has any Provider Role, then they are allowed to access additional service transactions**
  - If a user has exactly one Provider Role, then ECHO will use that Provider whenever a Provider Transaction is executed
  - If a user does not have exactly one Provider Role, then they must run the SetProviderContext method on the Session Manager to let ECHO know which Provider they are acting on the behalf of when they execute a Provider Transaction
    - Note: If a User has the Admin role, they can use SetProviderContext to become any Provider. If they have exactly one Provider Role in addition to the Admin Role, then they will not need to call SetProviderContext to act on behalf of that Provider. If they have more than one Provider Role, then they must call SetProviderContext

Provider Account Service

The Provider Account Service allows a registered data provider to manage their account information. This includes the ability to add, delete, present and update contacts, present provider information, change provider password, set and present provider policies.

Global Science & Technology, Inc.

Provider Contact Related Transactions

- **AddContact**
- **UpdateContact**
- **PresentContacts**
- **DeleteContact**

AddContact

- **Add one or more new contacts for a registered provider. A contact is an individual who has a named role within the organization. It has role name, first name, last name, address information, phone information and e-mail. Each contact has a unique role name like "shipping manager," or "accounts receivable".**
- **If a contact named "order manager" is created, then emails that are sent to users regarding orders will be CCed to this email address**
- **Caveats:**
  - Contact role must be unique for each provider, adding a contact with existing role name is not allowed. Error message "Contact Role: xxx ALREADY exists" if this is attempted.
  - Contact role name, first name and last name are required, error message will return if either one of them is not provided.
  - The whole transaction fails if any single piece of contact information violates the rules.

UpdateContact

- **Update the stored information of one or more registered provider contacts**
  - Updated information includes first name, last name, address information, phone information and e-mail information
  - The entire contact is replaced
- **Caveats:**
  - Updating a non-existing contact is not allowed, or it will return "Contact Role: XXX DOES NOT exist"
  - Role name, first name and last name should be filled, or it will return "Contact xxx is NULL"
  - The whole transaction will fail if any single piece of contact information violates the rules.

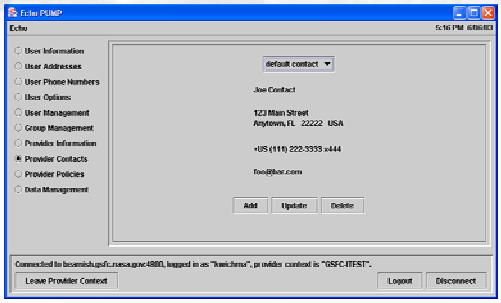
PresentContacts

- **Present one or more provider contacts' information, contact information, which includes the contact role, contact role's first name and last name, address information, phone information and e-mail information.**
- **Caveats:**
  - Presenting an non-existing contact is not allowed. It will return "Contact Role xxx is not found"
  - The whole transaction will fail if any single requested contact does not exist.

DeleteContact

- **Delete the specified contacts for this provider by specifying the contact role names.**
- **Caveats:**
  - Deleting a non-existing contact is not allowed. It will return "Contact Role: xxx DOES NOT exist"
  - The whole transaction will fail if any single requested contact does not exist.

PUMP Manage Contacts Screen



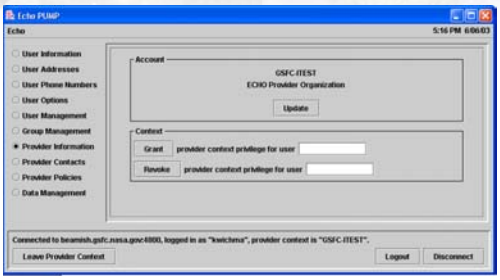
ProviderInformation

- **PresentProviderInfo**
  - Present the provider information and provider contacts information
  - Currently provider information only includes organization information
- **UpdateProviderInformation**
  - Currently just updates the organization information

Provider Role Functions

- **GrantProviderAccess**
  - Used by a registered user who has the provider role for some provider to grant the provider role for that provider to another registered user
- **RevokeProviderAccess**
  - Used by a registered user who has the provider role for some provider to revoke the provider role for that provider from another registered user who already has the provider role

PUMP Provider Information and Role Management Screen



A Well Deserved Break

Global Science & Technology, Inc.

ECHO Ingest and Metadata Update Process

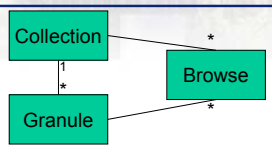
Global Science & Technology, Inc.

Echo Metadata Ingest

- **Metadata Preparation Requirements for Data Providers**
- **Metadata Ingest Configuration and Process**

As a clearinghouse of metadata, ECHO must provide an efficient mechanism for participating metadata providers to share their metadata with ECHO. ECHO has implemented this mechanism using an FTP interface to allow large files to be efficiently transferred to ECHO and to match existing provider distribution capabilities.

Simplified ECHO Metadata Model



- **At the highest level, there are three main entities that are part of the ECHO conceptual data model**
  - Collection – A grouping of granules typically based on a common source of the granules
  - Granule – The lowest level item retrievable from a provider that is uniquely described in ECHO
  - Browse – Some kind of binary or ASCII file used to provide a user with a quick view of the data. This could be a scaled down version of the imagery, or a histogram of the data, or some other representation

Metadata Management Through Ingest

- **What is currently managed through this interface?**
  - Collections, granules and browse
- **What are the lifecycle metadata management capabilities of ingest?**
  - Insert, update, delete
- **What is the granularity of access?**
  - Record level – The entire metadata record for a collection or granule is updated as a whole
- **How does the provider express the metadata?**
  - For collections and granules, an XML file that matches the DTD for collections or granule ingest files
  - For browse, both an XML file that matches the browse ingest file DTD and a set of binary browse files
- **How does the provider send the metadata to ECHO**
  - FTP – More details to follow

Metadata Preparation Requirements for Data Providers

- **Generating input XML files**
  - ECHO DTD compliant preferred
  - BMGT format – handled by an ingest proxy
- **Data Integrity and Presentation**
  - Discipline Keywords
  - Measured Parameters
  - Spatial Data
  - Date Format

Overview of ECHO Collection Metadata

- **Collection**
  - Basic information (ShortName, VersionID, description, processing level…)
  - Spatial (point, line, polygon, bounding box, circle)
  - Temporal (range, single shot, periodic)
  - Source information (platform, instrument, sensor)
  - PSA definitions
  - Keywords (discipline (GCMD stack), spatial, temporal)
  - Campaign
  - Science Review
  - Collection online resources and online access URLs
  - Algorithm package
  - Contact information
  - CSDT information
  - Other

Overview of ECHO Granule Metadata

- **Granule**
  - Basic information (GranuleUR, …)
  - Spatial (point, line, polygon, bounding box, circle)
  - Temporal (range, single shot, periodic)
  - Source information (platform, instrument, sensor)
  - PSA
  - Collection
  - Campaign
  - Science review
  - Online access URLs, granule online resources
  - PGE information
  - Measured parameters
  - Spatial domain Information
  - Associated granules (Input, Processing, History, QA….)
  - Other

Understanding the ingest DTDs

- **ECHO ingest DTDs**
  - Collection: http://www.echo.eos.nasa.gov/dtd/v5.0/collectioningest.dtd
  - Granule: http://www.echo.eos.nasa.gov/dtd/v5.0/granuleingest.dtd
  - Browse: (Currently insert only) http://www.echo.eos.nasa.gov/dtd/v5.0/browseingest.dtd
    - Browse ingest consists of an XML file matched with a set of binary files that are the browse
    - The XML file matches the GranuleUR (or Collection ID) against a set of file names
    - ECHO hosts the browse in an http server and provides a link within the granule's (or collection's) metadata to the URLs of the browse

Understanding the ingest DTDs (Cont)

- **Collections and granule XML files have four possible sections containing the payload of collection or granule metadata – insert, update, delete and list**
  - New – the payload of metadata is all to be inserted into the clearinghouse and that they do not already exist
  - Update – the payload of metadata is to be used to find existing records in the clearinghouse and change them to match the payload
  - Delete – the payload is a list of unique identifiers that are to be deleted from the system
  - List – the payload of metadata is to be synchronized with the clearinghouse (used by BMGT)
    - Provider delete date/timestamp is used to indicate deletion
- **Currently, inserts, updates and synchronizations all occur in the same way**
  - If a record already exists, it is updated; if it doesn't exist, it is created
  - The Provider Last Update Date/Timestamp is checked on metadata to be updated versus the metadata already in the clearinghouse
    - Only newer data can update the clearinghouse: this prevents synchronization issues from poisoning the clearinghouse

Data Dictionary

- **Data dictionary can be found at** http://www.echo.eos.nasa.gov/documents/ECHO_DTDTAG_DICTIONARY1.xls
  - The length of the field is represented as VA#### where the #### is the maximum number of characters
- **Some salient details:**
  - Date data – Format is yyyy-mm-dd hh:mm:ss unless otherwise specified
  - Keywords – GCMD Standard Compliant
    - No temporal keywords at this time maintained by GCMD
- **How to uniquely identify elements in ECHO - Assumptions**
  - Collection: Collection short name + version (obsolete in future)
    - Will be simply datasetID in the future
  - GranuleUR
  - Short name of platform, instrument, sensor
  - Browse Internal file name
  - Additional attributes name (PSA)
  - Provider order ID (future- to be resolved)
  - Other

Metadata Update Process

- **The Metadata Update capability allows certain fields to be updated independently of the rest of the granule**
  - In 5.0, this works for granule OnlineURL (for Data Pools), granule online resources and QA flags only
  - The ECHO last update timestamp is used to make sure the update is not performed out of sequence – i.e. if the ProviderLastUpdateDateTime is earlier than ECHO's copy, then the update is not performed
    - If it is performed, then the SaveDateTimeFlag indicates to ECHO whether to set its ProviderLastUpdateDateTime with the value provided in this message

Metadata Update DTD

```
<?xml version="1.0" encoding="UTF-8"?>
<!ELEMENT ProviderAccountService (UpdateMetadata)>
<!ELEMENT UpdateMetadata (Collection*, Granule*)>
<!ELEMENT Collection (Target+, (Add | Update | Delete)+)>
<!ELEMENT Granule (Target+, (Add | Update | Delete)+)>
<!-- Target+ allows the same change to be made to several different granules or collections
simultaneously.  This is especially useful for bulk deletions of OnlineURLs. -->
<!ELEMENT Target (ID, ProviderLastUpdateDateTime, SaveDateTimeFlag?)>
<!-- SaveDateTimeFlag is the flag that allows echo to update the last update date time for
Target. The default is SAVE -->
<!ELEMENT Add (QualifiedTag, MetadataValue)>
<!ELEMENT Update (QualifiedTag, MetadataValue)>
<!ELEMENT Delete (QualifiedTag+)>  <!ELEMENT QualifiedTag (#PCDATA)>
<!ELEMENT MetadataValue (#PCDATA)>
<!ELEMENT ProviderLastUpdateDateTime (#PCDATA)>
<!ELEMENT SaveDateTimeFlag (SAVE | DONTSAVE)>
<!ELEMENT SAVE EMPTY>  <!ELEMENT ID (#PCDATA)>
<!ELEMENT DONTSAVE EMPTY>
```


Metadata Update Example

```
<ProviderAccountService>
 <UpdateMetadata>
  <Granule>
   <Target>
    <ID>GranuleUR1234</ID>
    <ProviderLastUpdateDateTime>Dec 31 2002 12:00:00:000AM</ProviderLastUpdateDateTime>
    <SaveDateTimeFlag>
     <DONTSAVE/>
    </SaveDateTimeFlag>
   </Target>
   <Add>
    <QualifiedTag>OnlineAccessURLs/OnlineAccessURL/URL</QualifiedTag>
    <MetadataValue>http://www.echo.com</MetadataValue>
   </Add>
   <Add>

<QualifiedTag>OnlineAccessURLs/OnlineAccessURL[URL="http://www.echo.com"]/MimeType</QualifiedTag>
    <MetadataValue>hdf</MetadataValue>
   </Add>
  </Granule>
 </UpdateMetadata>
</ProviderAccountService>
```


Update Example	XPath	Value
Granule URL Delete One	OnlineAccessURLs/OnlineAccessURL[URL="old_url"]	NA
Granule URL Delete All	OnlineAccessURLs	NA
Granule URL Update	OnlineAccessURLs /OnlineAccessURL[URL="old_url"]	New url
Granule URL Insert	OnlineAccessURLs /OnlineAccessURL/URL	url
Granule Online resource Delete One	GranuleOnlineResources/OnlineResource[OnlineResourceURL="old_url"]	NA
Granule Online resource URL Update	GranuleOnlineResources/OnlineResource[OnlineResourceURL="old_url"]	New url
Granule Online resource type Update	GranuleOnlineResources/OnlineResource[OnlineResourceURL ="old_url"]/OnlineResourceType	New type
Granule URL mime type Insert	GranuleOnlineResources/OnlineResource[OnlineResourceURL ="old_url"]/OnlineResourceMimeType	Mime type
Granule Automatic QA Flag Update	MeasuredParameters/MeasuredParameter [ParameterName="p1"]/QAFlags/AutomaticQualityFlag	New flag
Granule Automatic QA Flag explanation Update	MeasuredParameters/MeasuredParameter [ParameterName="p1"]/QAFlags/AutomaticQualityFlagExplanation	New explanation
Granule Operational QA Flag Update	MeasuredParameters/MeasuredParameter [ParameterName="p1"]/QAFlags/OperationalQualityFlag	New flag
Granule Operational QA Flag explanation Update	MeasuredParameters/MeasuredParameter [ParameterName="p1"]/QAFlags/OperationalQualityFlagExplanation	New explanation
Granule Science QA Flag Update	MeasuredParameters/MeasuredParameter [ParameterName="p1"]/QAFlags/ScienceQualityFlag	New flag
Granule Science QA Flag explanation Update	MeasuredParameters/MeasuredParameter [ParameterName="p1"]/QAFlags/ScienceQualityFlagExplanation	New explanation

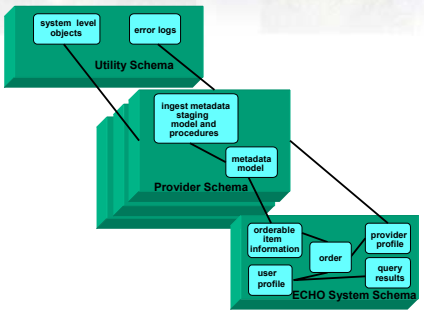
Describing URLs in ECHO

- **Type**
  - This is an ECHO controlled text field that is used to indicate what the URL is used for (Guide document, Order web page, etc.)
  - This is not present in OnlineAccessURLs because they are links to the data, so no type is necessary
- **MIME Type**
  - This is used to indicate what kind of document the client will find at the other end of the URL
  - It is intended to follow the internet's MIME standards, but could have extensions for dealing with non-official file types such as HDF-EOS
- **Description**
  - This is a free text description of the file being linked to

Configuring a Provider for Ingest

Global Science & Technology, Inc.

Database Design



ECHO Database Architecture

- **Provider database account and schema**
  - Table structure
  - Provider specifics
    - Date format process
    - PSA process
    - Order process
    - Spatial coordinate system

Provider Update Requirements

- **Metadata updates occur very soon after the Provider submits them (within 5 minutes)**
- **More frequent updates provide users with more up-to-date metadata to search**
- **Updates should match the rate of change of the data, but not to the point of overload**
- **ECHO staff should work with Provider to determine best fit for update schedule**
- **GOAL: ECHO should accurately represent Provider's metadata so that clients will be able to make accurate searches and orders**

Ingest Process

- **Schedule Ingest Process kick off**
  - Every 5 minutes
- **Ingest metadata for every data provider who is on a pre-configured list of data providers**
  - Check completeness of files in each data directory
  - Split big files into smaller files (≤1000 granules)
  - Decompose the XML files to multiple loadable ASCII data files
  - Load data into Oracle using SQL*Load
  - Filter and update operational tables
- **Performance**
  - Ingest: Approximately 10 minutes for about 1000 collection/granules

Ingest Performance

- **Ingest performance varies based on complexity of metadata, the degree the metadata is indexed in the clearinghouse, the amount of metadata in the clearinghouse (per provider), the operation performed (insert, update, delete) and the query load on the system at the time**
- **Inserts range from 6000-17000 operations per hour**
- **Ingest updates can be as low as 2000 operations per hour**

FTP Interface Configuration

- **Each provider has an FTP account**
  - Home directory is parent directory of where metadata should go
  - Password is not used, limited by IP
- **FTP directory structure**
  - data
    - collection
    - granule
    - browse
- **How ECHO recognizes completion of FTP transmission**
  - Catch the ending tag in the XML files
  - Verify browse XML file parameters (name and size) against browse image files

Ingest Processing Order

- **For each provider:**
  1. Ingest collection files
  2. Ingest granule files
  3. Metadata update files
  4. Ingest browse files

Error Handling

- **Data type inconsistency**
  - Caught by SQL*Load
    - The entire collection or granule will be ignored. The error information will be recorded and reported to the data provider
- **Unrecognized TAGs and inconsistent XML format**
  - Caught by XML file validation
    - The XML input file will not be processed. The error information will be recorded and sent to the data provider.
- **Duplicated entries, data outside table constraints, and invalid spatial data**
  - Caught by updating procedure. The collection/granule is ignored. The error information is recorded and is sent to the data provider.
- **Granule associated with a collection that has not been registered in ECHO yet**
  - Granule will be rejected. Error information is recorded and sent to the provider.
- **Out of date**
  - The provider last update date/timestamp indicates that the "new" data is older than the existing data in the clearinghouse. The collection/granule is ignored.
- **Collection/Granule deletion of non-existent data**
  - A deletion request comes in for data that does not exist. The request is ignored.

The Ingest Summary Report

- **Contents**
  - Lists files received (name and size)
  - Number of records processed
  - Number of records inserted
  - Number of records replaced
  - Number of records deleted
  - Number of records out of date
  - Number of records duplicated
  - Number of collections/granules with invalid spatial data
  - Number of records ignored due to data errors
  - Number of records without an associated collection
  - Number of records that were attempted to be deleted, but didn't exist
  - Error message(s) if any
  - Metadata update detailed error messages
  - Anything else?
- **Sent to the data provider specified email address**

Ingest File Error Summary Report

- **Lists of files that are not XML files**
  - No XML header
  - Don't contain XML
- **List of files that are not in the correct ingest directory**
- **List of files that did not validate against the DTD**
- **Report is sent to the data provider via email**

Representing Spatial Data for Ingest

Global Science & Technology, Inc.

About Spatial Data

- **ECHO leverages Oracle's spatial capabilities to perform spatial search and to represent provider metadata**
- **Both collections and granules can be spatially represented**
  - If a collection or granule represents the entire Earth, there is no need to spatially index it
  - ECHO is asking Providers to set a spatial keyword to "Global" in order to indicate this condition
    - Version 5.5 will have a special Global flag for the provider to set
- **Oracle 9i version 2**

Understanding Oracle Spatial

- **Coordinate system**
  - Standard Coordinate System Definition (OGC)
    - ECHO uses Cartesian and Geodetic models
  - User Defined Coordinate System
    - Not used by ECHO
- **Data type**
- **Data presentation**
- **Spatial Indexing**
  - Quad tree Indexing
    - Not used by ECHO
  - R-tree Indexing
    - ECHO uses R-tree Indexing

Cartesian vs. Geodetic Models

- **Cartesian Model**
  - The Earth is represented by a rectangle (-180, -90, 180, 90)
  - Spatial objects cannot cross the International Date Line
  - Spatial objects cannot cross the poles
  - Polygons can be as large as the whole Earth
- **Geodetic Model**
  - The Earth is represented by a spheroid
  - Spatial objects may cross the International Date Line
  - Spatial objects may cross the poles
  - No single polygon may have an area equal to or larger than half of the Earth's area

Oracle Spatial Data Types

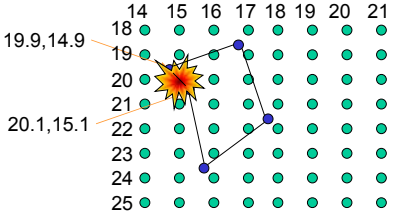
- **Point and Multiple points**
- **Line**
  - Straight line (two points connected by a great circle)
  - Arc (two points connected by a user specified arc)
  - Multiple lines (including straight line and/or arc)
- **Circle**
  - Three points on the circumference
- **Polygon**
  - Polygon
  - Polygon with hole
  - Polygon with multiple holes
  - Multiple polygons (including all of above polygons)
- **Bounding Box**
  - Only valid for the Cartesian model
- **Combination of any of the above**
  - The combination is referred to as a single spatial coverage

Oracle Spatial Data Presentation

- **Polygon**
  - Outer ring
  - Inner ring - this represents a hole
  - Polygons crossing International Date Line or the pole are not supported in the Cartesian model
  - No polygon may have an area larger than or equal to half of the area of the Earth in the Geodetic model
- **In general**
  - In polygon, vertices must be distinct
  - No overlap between any two objects in one spatial record
- **Spatial Data validation - Spatial resolution**
  - The lower the value of spatial resolution, the bigger the spatial index will have to be in ECHO
  - This should be set to a reasonable value

Understanding Spatial Resolution

- **Latitude Resolution = 0.5, Longitude Resolution = 0.5**
  - The two blue dots will be the same point
- **Latitude Resolution = 0.1, Longitude Resolution = 0.1**
  - The two blue dots will be different points



How ECHO applies Oracle Spatial

- **Coordinate systems supported**
  - Flat Cartesian system
  - Geodetic system
- **Data types supported**
  - Point and multiple points
  - Straight line and multiple lines
  - GPolygon(s) - connected with straight line (great circle if in geodetic model)
    - Only allows one hole
  - Polygon(s) - connected with straight line (great circle if in geodetic model)
    - Allows holes
  - Bounding Box(es)
  - Combinations of these
- **Indexing - R-tree indexing (incremental)**

What is Expected from Data Providers

- **Select appropriate coordinate system for your spatial data**
- **Select appropriate data type for your spatial data**
  - Point
  - Line
  - Polygon
  - Multi polygon
  - Circle (input with center point lat/lon and radians in meters)
  - Bounding box
    - This is only appropriate for the Cartesian model
    - Future: ECHO will be able to convert bounding box into polygon for use in the Geodetic model
- **Spatial information must provide a reasonable approximation of the footprint of the data**
  - Spatial searches are only as good as the data provided
  - Densification may be necessary to connect points as expected

Geodetic View of Sample Polygon

Cartesian View of Sample Polygon

Cartesian Model Providers

- **Apply Flat Cartesian coordinate system**
  - For GPolygon type, put vertices in a clockwise order
  - For Polygon type, put vertices in a clockwise order for both inner ring and outer ring
  - Split the GPolygon if it crosses the International Date Line or poles
  - Circle data as defined in the DTD (center point, radius) will be translated into a polygon using Oracle spatial utility

Geodetic Data Providers

- **Apply Geodetic coordinate system**
  - Geodetic
    - GEOGCS ["Longitude / Latitude (WGS 84)", DATUM ["WGS 84", SPHEROID ["WGS 84", 6378137.000000, 298.257224]], PRIMEM ["Greenwich", 0.000000], UNIT ["Decimal Degree", 0.0174532925199433O]]
  - Prepare data correctly for the Geodetic system
    - For Polygon type, put vertices in a clockwise order for both inner ring and outer ring
    - Provide adequate density of the vertices
      - ECHO assumes that points are connected via the shortest distance between the points
    - Polygon coverage limitation – no more than half of earth
    - Bounding box conversion (future)

Lunch

Provider and ECHO Order Interactions

ECHO brokers orders from clients to Providers with the expectation that the Provider will send the data directly to the client. This section covers the transactions both from ECHO to the Provider and from the Provider to ECHO concerning quotes, orders and cancellations.

Components of An Order Sent to Providers

- **An order consists of several pieces of information**
  - Order ID
  - Provider Tracking ID
  - Shipping, Billing and Contact Addresses
  - A list of catalog items to be ordered that belongs to the provider
  - Quantity of the items
  - Options of the items

Options in Order

- **Options are used here to represent the Provider defined settings that are needed to completely describe an order but are unique to a provider**
- **Order Entry Service allows for options at 3 levels**
  - Order - Options that apply across the entire order
    - Currently not used
  - Provider Order - Options that apply to every item being sent to a Provider
    - Currently not used, but could be
  - Line Item - Options that apply only to the smallest unit being ordered
    - Package options are specified here (media, shipping, etc.)
- **Also used in other places, but may be called policies or preferences.**

ECHO Option Framework Explained

The Option Framework is used in a number of places throughout ECHO to provide a mechanism for introducing new data structures into the system during operations.

Global Science & Technology, Inc.

What are Options?

- **ECHO uses a common infrastructure for allowing abstract, definable items to be specified at run-time**
  - If the structure of something is not known when ECHO is developed, then the option infrastructure can support representing that structure without changes to the ECHO code
  - The options infrastructure allows a client to see what the template (Option Definitions) for a structure is, as well as any previous selections
  - It also allows a client to set (select) values in the structure according to the rules in the Option Definitions Template

Options

- **Option Definitions - describe the options available for a given item. Some options may be declared as required.**
- **Option Selections - the actual value associated with an option definition for a particular item.**

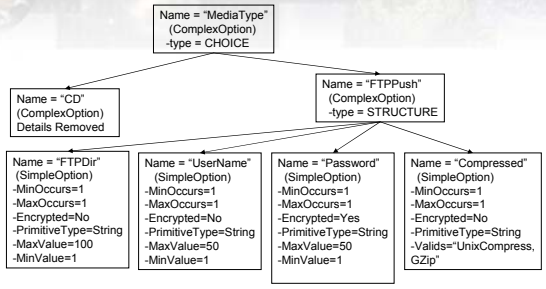
Option Definitions

- **Simple Option Attributes**
  - Option Name- The string that is used to refer to this node of the option tree
  - MinOccurs - The minimum number of times this node should appear in the option tree
  - MaxOccurs - The maximum number of times this node should appear in the option tree
  - Encrypted - When representing this field for display or input, use a text box that hides what is being typed (password box)
  - Primitive Type - The type of the option being represented
    - String, Integer, Boolean, Date, Double
  - Valids - If a string, an enumeration of valid values, if a number, an enumeration of valid values
  - Min Value - If a number, the minimum value, or if a string, the minimum length
  - Max Value - If a number, the maximum value, or if a string, the maximum length
- **Complex Option Attributes**
  - Structure - All contained options should be present in the Option Selection
  - Choice - Only 1 of the contained complex options should be present in the Option Selection

Option Selections

- **The Option Selection tree should parallel the Option Definitions tree**
- **Simple Option Attributes**
  - Option Name – Should equal Option Definition name at this node in the tree
  - Value – The value that the user wants assigned
    - This will be validated against the information in the Option Definition for this node
- **Complex Option Attributes**
  - Option Name – Should equal Option Definition name at this node in the tree
  - ComplexValue – This is a container for placing the Option Selections for the corresponding child Options
- **Option Selections are validated against the Option Definition when Options are set**

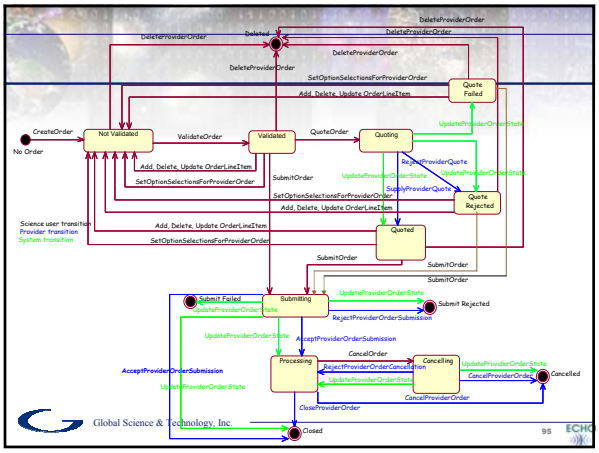
Options Definition Illustration Example




```
<OptionSelection>
    <ComplexOptionSelection>          complex option consists of
                                      an option name and a complex value
      <OptionName>ECS-TEST PKG1</OptionName>
      <ComplexValue>                  the complex value consists of
                                      zero or more simple option
                                      and zero or more complex option
        <SimpleOptionSelection>                    simple option consists of
          <OptionName>Production Option</OptionName>   an option name
            <Value>Native Granule</Value>              and one or more value
        </SimpleOptionSelection>
        <ComplexOptionSelection>
          <OptionName>Media Type</OptionName>
          <ComplexValue>
            <ComplexOptionSelection>
              <OptionName>CDROM</OptionName>
              <ComplexValue>
                <SimpleOptionSelection>
                  <OptionName>Compression</OptionName>
                  <Value>GZip</Value>
                </SimpleOptionSelection>
                <SimpleOptionSelection>
                  <OptionName>DDISTMEDIAFMT</OptionName>
                  <Value>Rockridge</Value>
                </SimpleOptionSelection>
              </ComplexValue>
            </ComplexOptionSelection>
          </ComplexValue>
        </ComplexOptionSelection>
      </ComplexValue>
    </ComplexOptionSelection>
</OptionSelection>
```

States of a Provider Order

Open Order States	Closed Order States
• Not Validated	• Deleted
• Validated	• Submit Failed
• Quoting	• Submit Rejected
• Quoted	• Cancelled
• Quote Failed	• Closed
• Quote Rejected	
• Submitting	
• Processing	
• Cancelling	



How Does ECHO Communicate with Providers?

- **User actions on orders that trigger communication with providers:**
  - Quote, Submit, Cancel
- **Two ways to handle these requests**
  - Synchronous Action (i.e. ECHO waits for a response)
  - Asynchronous Action (i.e. contact ECHO with a response)
- **Retry Mechanism**
  - If no response received from provider, ECHO will issue another request after waiting for a period of time
  - Provider can set policy to determine the maximum numbers of attempts for retry and waiting time between retries
- **TrackingID**
  - ECHO provides a TrackingID on its first communication to the Provider
  - The Provider has the option of replacing it with its own TrackingID
  - ECHO will use the "current" TrackingID for all subsequent communications with the Provider and make that available to the client

What providers give back to users?

- **Status Message**
  - Timestamped message records the status history of the provider order that can be viewed by the user using PresentProviderOrder or PresentOrder transaction in Order Entry Service
- **Additional information for submit response**
  - Order Information
    - total price for the order (including all prices below)
      - This is the only mandatory item
    - price for the data
    - price for the media that stores the data
    - price for shipping
    - price for handling
    - any discount made to the total price
    - quantity of media used to store the data
    - estimated ship date for the order
      - Assuming the order is placed immediately
    - latest date for cancellation of the order
    - additional information
      - This is a place for providers to put free text

What providers give back to users?

- **Additional information for quote response**
  - Provider Quote
    - total price for the order (includes all prices below)
      - Mandatory
    - price for the data
    - price for the media that stores the data
    - price for shipping
    - price for handling
    - any discount made to the total price
    - recommended media to store the data
    - quantity of media used to store the data
    - expiration date for the quote
      - Mandatory
    - additional information

Sample XML ECHO→Provider

```
<SubmitProviderOrderRequest>
  <Order>
      <ProviderID>ORNL</ProviderID>
      <ECHOOrderID>ORNL1234</ECHOOrderID>
      <ProviderTrackingID>ORNL1234</ProviderTrackingID>
      <UserInformation>
            <ECHOUserID>…</ECHOUserID>
            <ShippingAddress>…</ShippingAddress>
            <BillingAddress>…</ BillingAddress>
            <ContactAddress>…</ContactAddress>
            <NetworkAddress>…</NetworkAddress>
      </UserInformation>
      <OrderLineItem>
            <CatalogItemID>9876</CatalogItemID>
            <quantity>2</quantity>
            <OptionSelection>…</OptionSelection>
      </OrderLineItem>
            …..
  </Order>
</SubmitProviderOrderRequest>
```


Synchronous Action

- **User sends a request to ECHO to quote, submit, or cancel a provider order.**
- **ECHO sends that message to a specified provider location using the communication protocols defined.**
- **ECHO waits for provider's response which will come after the provider performs the requested action.**
  - Provider agrees to TrackingID as discussed on previous slide.
  - The provider reacts to the request immediately and sends back an answer as part of that response.

Synchronous Sample Response

```
<SubmitProviderOrderResponse>
  <Order>
      <ECHOOrderID>1234</ECHOOrderID>
      <ProviderTrackingID>ORNL:1234</ProviderTrackingID>
      <ProviderOrderAcceptance>
            <latestCancelDate>01-01-2001 13:42:00</latestCancelDate>
            <estimatedShipDate>01-02-2001</estimatedShipDate>
            <additionalInformation>ID5 will ship 1/4</additionalInformation>
            <TotalPrice>…</TotalPrice>
            <ShippingCost>…</ShippingCost>
            <MediaCost>…</MediaCost>
            <HandlingCost>…</HandlingCost>
            <DataCost>…</DataCost>
            <Discount>…</Discount>
      </ProviderOrderAcceptance>
      <StatusMessage>…</StatusMessage>
  </Order>
</SubmitProviderOrderResponse>
```


Asynchronous Action

- **User sends a request to ECHO to quote, submit, or cancel a provider order.**
- **ECHO sends that message to a specified provider location.**
- **ECHO receives acknowledgement that the message was received.**
  - Provider agrees to TrackingID as discussed on previous slide.
  - The provider schedules the action and will later initiate a message to ECHO with the result.

Asynch. Sample Response

```
<SubmitProviderOrderResponse>
  <Order>
    <ECHOOrderID>1234</ECHOOrderID>
    <ProviderTrackingID>ORNL:1234</ProviderTrackingID>
  </Order>
<SubmitProviderOrderResponse>
```


Provider Order Management Service

This service enables a Provider to give ECHO updates on Orders, cancel them, check their order history and make asynchronous responses to ECHO's requests

Provider Order Management Service (POMS)

- **Allows the providers to take action (either accept or reject) on the user's request (submit, quote, cancel) if asynchronous action was used**
- **Allows the providers to update the order status**
  - Update the status message
  - Close an order in process
  - Cancel an order for any reason
- **Allows the providers to manage order more efficiently**
  - View order histories of open/closed order
  - Change tracking ID

Submit Transactions

- **SubmitOrder (Order Entry Service)**
  - User requests for order processing.
- **AcceptProviderQrderSubmission**
  - Allows a provider to asynchronously accept a user's order processing request
  - Provider supplies order information including total price, shipping/handling, estimated ship date, latest cancel date, etc.
  - Provider can choose manual status updates, or go ahead and close the order (no further updates are guaranteed)
  - ECHO updates provider order state from SUBMITTING to PROCESSING (if provider choose to manually update order status), or CLOSED (if provider choose to not guarantee update to any order status, order will be closed upon completion of processing).
  - ECHO appends status message to the provider order and notifies user if needed.
- **RejectProviderQrderSubmission**
  - Allows a provider to asynchronously reject a user's request to quote a particular order.
  - Provider gives a reason for rejection
  - ECHO updates provider order state from SUBMITTING to SUBMIT_REJECTED
  - ECHO appends status message to the provider order and notifies user if needed.

AcceptProviderOrderSubmissionRequest Sample

```
<ProviderOrderManagementService>
  <AcceptProviderOrderSubmissionRequest>
    <ProviderTrackingID>ORNL:1234</ProviderTrackingID>
    <OrderID>1234</OrderID>
    <SubmissionInformation>
      <totalPrice>10.00</totalPrice>
      <latestCancelDate>11/10/02</latestCancelDate>
      <estimatedShipDate>11/20/02</estimatedShipDate>
      <dataPrice>5.00</dataPrice>
      <mediaPrice>1.00</mediaPrice>
      <shipping>3.00</shipping>
      <handling>1.00</handling>
      <discount>0</discount>
      <quantityOfMedia>1</quantityOfMedia>
      <additionalInformation>Holiday on 11/15/02, office will be closed</additionalInformation>
    </SubmissionInformation>
    <StatusMessage>The order is accepted and in process, check frequently for status</StatusMessage>
    <UpdateMechanism><MANUAL/></UpdateMechanism>
  </AcceptProviderOrderSubmissionRequest>
</ProviderOrderManagmentService>
```


RejectProviderOrderSubmissionRequest

```
<ProviderOrderManagementService>
  <RejectProviderOrderSubmissionRequest>
    <ProviderTrackingID>ORNL:1234</ProviderTrackingID>
    <OrderID>1234</OrderID>
   <StatusMessage>The order is rejected because …</StatusMessage>
  </RejectProviderOrderSubmissionRequest>
</ProviderOrderManagmentService>
```

Quote Transactions

- **QuoteOrder (Order Entry Service)**
  - User requests quote on a specific order.
- **SupplyProviderQuote**
  - Allows a provider to asynchronously accept a user's request to quote a particular order.
  - Provider supplies quote information including total price, shipping/handling, quote expiration date, etc.
  - ECHO updates provider order state from VALIDATED to QUOTED.
  - ECHO appends status message to the provider order and notifies user if needed.
- **RejectProviderQuote**
  - Allows a provider to asynchronously reject a user's request to quote a particular order.
  - Provider gives a reason for rejection
  - ECHO updates provider order state from VALIDATED to QUOTE_REJECTED
  - ECHO appends status message to the provider order and notifies user if needed.

Cancel Transactions

- **CancelOrder (User Account Service)**
  - User requests that a provider order be cancelled.
- **CancelProviderOrder**
  - Allows a provider to asynchronously accept a cancel request made for a particular order.
  - ECHO updates provider order state from CANCELLING to CANCELLED.
  - ECHO appends status message to the provider order and notifies user if needed.
  - Provider can initiate this transaction even if the user has not requested a cancellation.
- **RejectProviderOrderCancellation**
  - Allows a provider to asynchronously reject a cancel request made for a particular order.
  - Provider gives a reason for rejection
  - ECHO updates provider order state from CANCELLING to PROCESSING
  - ECHO appends status message to the provider order and notifies user if needed.
  - The order is flagged, so that the user can not issue another cancel request

Process Order

- **Several transactions are useful if the provider chooses to update the order status manually when the provider accepts the order.**
  - CancelProviderOrder
    - Allows a provider to cancel an order in process for any reason, even if the user has not requested a cancellation
    - Provider gives a reason for cancellation
    - ECHO updates the provider order state from PROCESSING to CANCELLED
    - ECHO appends status message to the provider order and notifies user if needed.
  - CloseProviderOrder
    - Allows provider to close an order upon completion
    - ECHO updates the provider order state from PROCESSING to CLOSED
    - ECHO appends status message to the provider order and notifies user if needed.
  - UpdateStatusMessage
    - Provider can send updated message to the user, this is useful when the order is partially fulfilled
    - ECHO appends status message to the provider order and notifies user if needed.

View Order History

- **PresentOpenOrder**
  - Lists detailed information about all the orders who are in one of the following states:
    PROCESSING, CANCELLING
  - Optionally, the user can filter the open orders by specifying the order state
- **PresentClosedOrder**
  - Lists Detailed information about all the orders who are in the one of the following state:
    SUBMIT_FAILED, SUBMIT_REJECTED, CANCELLED, CLOSED
  - Optionally, the user can filter the closed orders by specifying the order state
- **PresentOpenOrderSummary**
  - Lists only the order ID and order state of open orders
- **PresentClosedOrderSummary**
  - List only the order ID and order state of closed orders

Change Tracking ID

- **ChangeTrackingID**
  - Provider accepts or replaces the tracking ID at its first communication with ECHO in response to either the submit or quote requests.
  - In some cases, the provider wants to change it in its own system.
    - This transaction is used to register the change to ECHO.

Provider Policies

Provider Policies

ECHO uses a generic policy/option mechanism to provide a run-time way of changing values that the business object operates on. This section describes this mechanism and how it applies to Providers.

Provider Policy Related Transactions

- **PresentPolicyDefinitions: Present a listing of the Policy Definitions that are currently available for this registered provider.**
  - The providers set the policy selection based on the policy definitions defined.
  - E.g. in the option definition sample, option "quote_retry_attempts" is defined as an integer type, minimum value is 1, maximum value is 1000, so when provider sets policy selection, it must be an integer between 1 and 1000.
- **SetPolicySelections: Set the policy selection for a registered provider based on the policy definition defined for this selection**
  - The selection will be validated according to the Policy Definition
- **PresentPolicySelections: Present a detailed listing of the policies that are currently set for this registered provider.**
  - A provider may request desired policy options by specifying the policy name
  - If policy name is not specified, the provider will receive a complete list of the current Policy Selections

Option Definition Sample

```
<ProviderAccountService>
    <PresentPolicyDefinitionsResponse>
        <OptionSelection>
            <SimpleOption>
                <Name>quote_retry_attempts</Name>
                <OptionCategory>
                Communication Setting
                </OptionCategory>
                <Description>
                the max attempts of quote retry
                </Description>
                <MinOccurs>1</MinOccurs>
                <MaxOccurs>1</MaxOccurs>
                <Encrypted>False</Encrypted>
                <PrimitiveTypeName>
                    <Integer/>
                </PrimitiveTypeName>
                <MinValue>1</MinValue>
                <MaxValue>1000</MaxValue>
            </SimpleOption>
        </OptionSelection>
    </PresentPolicyDefinitionsResponse>
</ProviderAccountService>
```

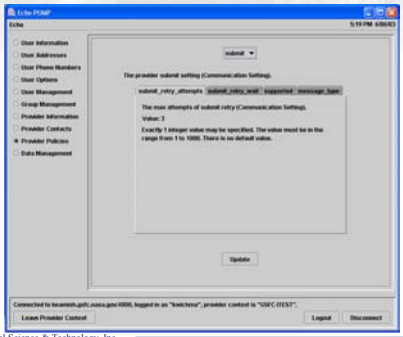

Option Selection Sample

```
<ProviderAccountService>
    <PresentPolicySelectionsResponse>
        <OptionSelection>
            <SimpleOptionSelection>
                <OptionName>
                max_attempts_to_quote_order
                </OptionName>
                <Value>4</Value>
            </SimpleOptionSelection>
        </OptionSelection>
    </PresentPolicySelectionsResponse>
</ProviderAccountService>
```


Provider Policies List

- **For each of submit, quote and cancel**
  - Supported
    - Does the provider support this transaction
  - Retry Attempts
    - Number of times to retry a message
  - Retry Wait
    - Number of seconds to wait between retries
  - Message Type
    - ECS/V0 ODL socket connection
      - Machine, Port
    - SOAP XML connection
      - Target method, target object, URI

PUMP Provider Policy Screen



Break Time

Global Science & Technology, Inc.

Access Control Lists

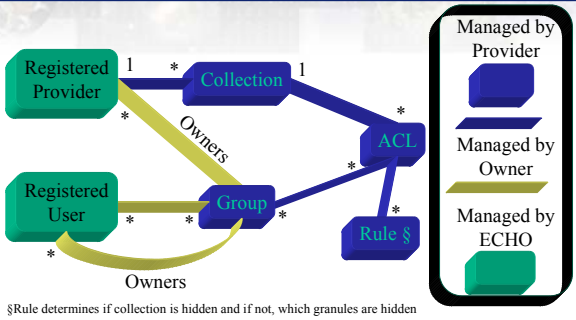
Protecting visibilty of metadata and limiting orders

Global Science & Technology, Inc.

Why do we have?

- **Originally ECHO was intended to be a clearinghouse of public metadata**
  - ECHO can now also serve as a clearinghouse for restricted (private) metadata
- **Desire in provider community to use a single system to represent their metadata**
- **Also a desire to leverage user interfaces developed for ECHO for those who have access to restricted (private) metadata.**

Metadata Visibility - Conceptual Model



§Rule determines if collection is hidden and if not, which granules are hidden

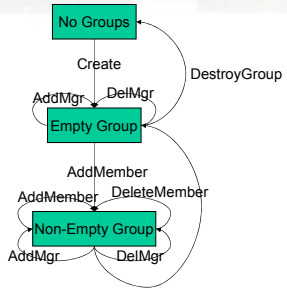
Key Concepts

- **Managers**
  - A list of users and/or providers
  - Can add or remove other managers
  - Can add or remove members of the group
  - Can contact the members or managers
- **Members**
  - A list of users and/or providers
  - Can contact other members
- **Real use: Metadata Visibility Permissions**

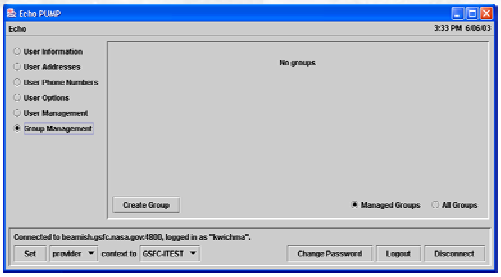
Why do we have groups?

- **Mechanism for aggregating users**
- **Communication**
- **Metadata Visibility**

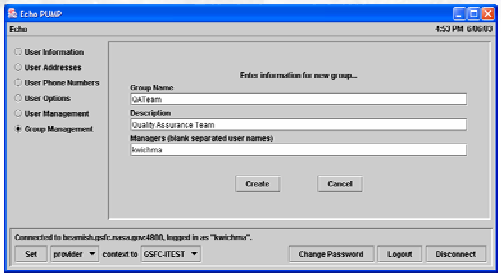
Group State Machine



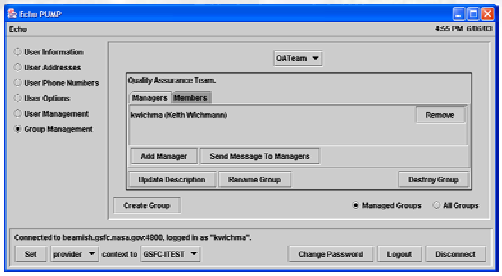
PUMP Group Management Functions Screen



PUMP Create Group Screen



PUMP Group Display - Managers Screen



PUMP Group Display - Members Screen

PUMP Group Management - Add Member Screen

Group Caveats

- **There can be multiple managers**
- **Any manager can add or remove other managers**
- **Managers are not automatically members of the group**
- **An owner has no control over the group, it is simply an attribute of the group that stores who originally created the group**
  - The owner will also be the first manager of a group if the group was created without specifying a manager

Data Management Service

- **Used by providers to manage access control to their metadata and ordering processes**
- **Composed of Rules, Conditions and Groups**
- **Rules and Conditions combined are used to specify subsets of a collection to which to apply access control**

ACL Approach

- **When do changes to ACLs take effect?**
  - In short, immediately (no batch updates needed)
  - The system is designed around the concept of provider control, and a provider's update is honored immediately in all ECHO transactions
- **Access Control Lists are honored at time of execution allowing a provider to make updates that will apply to existing result sets and orders**
  - This allows the provider to introduce new restrictions, remove existing restrictions, or arrange for a group to be updated
- **Certain types of restrictions are time-based and will expire automatically**
  - An ACL will allow access to a previously restricted result set when a temporal restriction expires
  - Does not require a provider to take an action

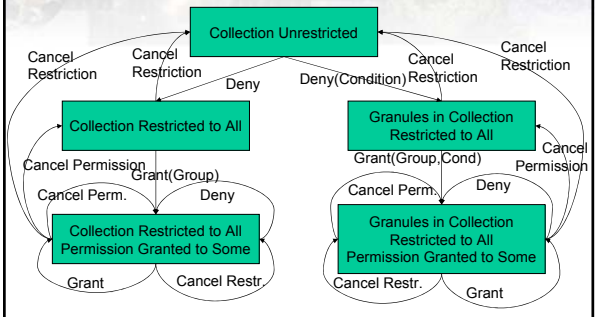
What access is being controlled?

- **Viewing**
  - Catalog Service and Subscription Service will prevent the presentation of controlled metadata
  - CAVEAT: Once a copy of the metadata is extracted from ECHO, it can be shared at will circumventing ECHO's access control functions exactly as is possible currently by ordering provider data and sharing with others
  - CAVEAT: Browse URLs that are part of the ECHO metadata are controlled in the sense that they are not presented if the metadata is not visible, but access is not checked for every client that accesses the URL directly
- **Ordering**
  - Order Entry Service prevents
    - Creation of orders with a restricted item
    - Adding restricted items to an order
    - Quoting orders with a restricted item
    - Submitting an order with a restricted item
- **Browse**
  - Browse has been discussed as being a separately controlled item, but is currently considered too complex

What is the granularity of control?

- **Collections**
  - Metadata Visibility: A collection's entire metadata description as well as all contained granules
    - CAVEAT: In the future, we plan to expose the name of the collection regardless of its access control state
  - Order: A collection or the granules within the collection
- **Granules**
  - An individual granule can be controlled by naming its Provider GranuleUR
- **Subsets of Granules within a Collection**
  - Frequently it is useful to specify groups of granules within a collection by some common attribute to protect
  - Subsets can be performed based on Time of Acquisition
    - Fixed time periods or rolling time periods can be used
    - Future: Subsets can be performed based on a provider controlled access flag, which can be set according to quality concerns, spatial location, etc.

DMS State Machine



Data Management Key Components

- **Conditions**
  - Temporal
  - Rolling Temporal
  - Boolean
- **Rules**
  - Restrictions
  - Permissions
- **Groups**
  - Only used with Permissions

Conditions – What are They?

- **Phraselets (unable to be evaluated)**
  - Temporal
    - Start: Sept 1, 2002
    - Stop: Sept 30, 2002
    - Temporal Target Field: Production Date, Acquisition Date, Provider Insert Date (5.0.1)
  - Rolling Temporal
    - 30 days
    - 10 years
    - Temporal Target Field: Production Date, Acquisition Date, Provider Insert Date (5.0.1)
  - Boolean
    - True
    - False

Evaluating Conditions - Comparators

- **Comparators + Conditions = Evaluation Mechanism**
- **Examples of Comparators**
  - <, >, <=, >=, ==, !=
- **How they are used**
  - Comparator: ==
  - Temporal Condition: start Sept 1, 2002; stop Sept 30, 2002
  - Result: The time frame between September 1, 2002 and Sept 30, 2002
- **Questions**
  - How would this evaluated with a != comparator?
  - Does a < comparator make sense in this context?

Rules

- **Provider-Only access**
- **References**
  - Condition and Comparator
  - Provider
  - Granule or Collection
  - Group (optional)
- **Types of Rules**
  - Restriction (does not use a Group)
  - Permission (requires a Group)

Restrictions

- **A form of a rule**
- **Prevents access to metadata**
- **Applies to all ECHO users**
- **Specific to an Action**
  - Actions: View or Order
- **Specific to a piece of metadata**
  - Collection MOD01
  - Granule SC:1234

Permissions

- **A form of a rule**
- **Enables access to metadata**
- **Applies to a group previously created**
- **Specific to an Action**
  - Actions: View or Order
- **Specific to a piece of metadata**
  - Collection MOD01
  - Granule SC:1234

Restriction Example

- **Desired Behavior: A provider wishes to prevent viewing of all "young" collections.**
- **Condition Information**
  - ConditionType: Rolling Temporal
  - Duration: 30 Days
- **Rule Information**
  - RuleType: Restriction
  - Comparator: Less Than
  - DataType: Collection
  - DataValue: "ALL"

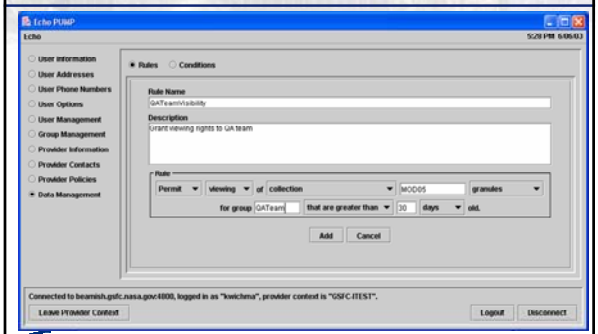
Permission Example

- **Desired Behavior: A provider wishes to grant viewing access to their internal testers**
- **Condition Information**
  - Condition Type: Boolean
  - Value: True
- **Rule Information**
  - RuleType: Permission
  - Comparator: Equals
  - DataType: Collection
  - DataValue: "ALL"
  - Group: Internal Testers

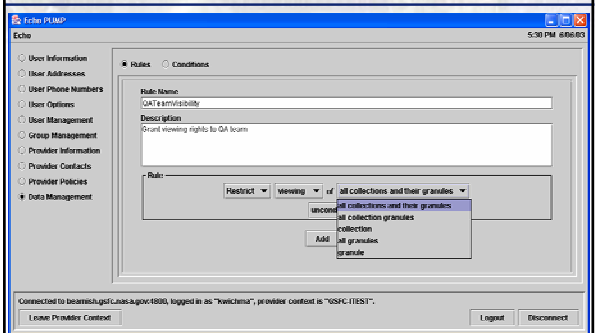
Data Management (ACL) - Provider Context

Create a Restriction - ACL - Provider Context

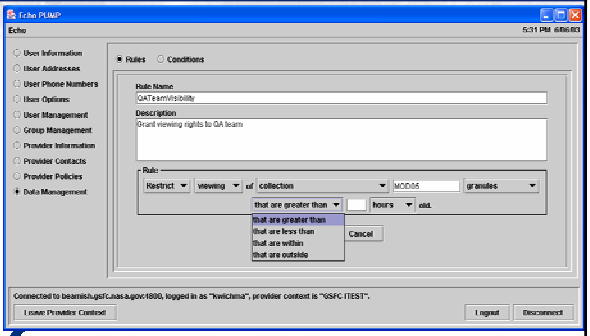
Create a Permission - ACL - Provider Context



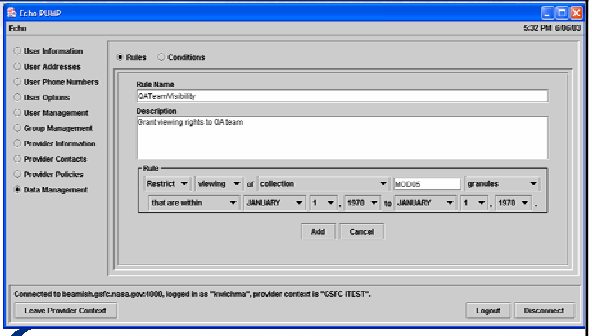
Restriction Options - ACL - Provider Context



Condition Options – ACL – Provider Context



Fixed Temporal Range Restriction Example



Operational Use Case

- **A new collection has been introduced into the provider's system. The provider wants to expose this data through ECHO, but there are some granules that are test and should not be publicly available.**
  - The new collection is registered in ECHO. The ECHO Ops group coordinates its introduction into the clearinghouse.
  - Immediately after it is registered, the provider uses PUMP to create a rule that blocks all access to that collection (Boolean condition on the collection). The provider also creates a permission to the correct group (Boolean condition on the collection).
  - Once the collection is ready for public display, a new restriction is created on the fixed time range in which the test data exists and in which no real data exists.
  - The first restriction is removed, allowing open access to the metadata by all ECHO users, with the exception of the test data. The test data can still be accessed by the groups that were given permission earlier.

Visibility "Gotcha's"

- **Rules evaluated optimistically**
  - Permissions supercede Restrictions
- **Rules short-circuit**
  - "ALL" DataValue keyword short-circuits individual DataValues
  - As soon as permission is located, visibility is granted
- **Conditions are re-usable**
- **Group management an external process**
- **It is possible to use the API to create ACLs that make no sense**

ECHO Operations Group (ECHO Ops) Support for Data Partners

Chao-Hsi Chang
Global Science and Technology
c-hchang@gst.com

Global Science & Technology, Inc.

ECHO PARTNER

What is an ECHO Data Partner?

Data Partners participate in ECHO by

- **Making metadata that represent their Earth Science data holdings available for search;**
- **Providing a mechanism to allow Client applications to access their data holdings, either by**
  - Order distribution
  - Online access

Overview of ECHO Ops Support for Data Partners

- **Data Partner Application and Setup**
- **Acclimation and Test Support**
- **Ingest Operations**
- **Partner Relations Management**

Data Partner Application and Setup

- **Data Partner Application**
  - Partner completes application form
  - Reviewed by ECHO Ops in conjunction with Project Managers to determine compatibility with ECHO scope and goals set forth by the ECHO Technical Committee (ETC)
- **Initial Setup**
  - Establish primary user accounts on operational and test systems
  - Provide access information and tools (e.g. PUMP)
  - Help to establish/implement policies, options, and other configurations in ECHO system
- **Establish an Operations Agreement (OA)**
  - Based on ESDIS ECHO template OA (currently being finalized)

Acclimation and Test Support

- **API Support**
  - Assist Partner in understanding relevant components of the ECHO API
- **Metadata Mapping**
  - Assist in the creation of metadata schema (if needed)
  - Assist in the creation of metadata ingest files that conform with the ECHO DTD
- **Access Control**
  - Assist Partner in using PUMP to manage data set access rules
- **Test Support**
  - Coordinate ingest of metadata on ECHO test system
  - Provide resources for Partner testing of search and order distribution processes
    - "Dummy" Providers
    - Canned XML scripts

Ingest Operations

- **Ingest Plan and Schedule**
  - Assist Partner in establishing ingest priorities
    - Among the different datasets or "Collections"
    - For different time periods in the historical archive
    - Requirements for keeping ECHO up to date with Partner's current and future data processing
  - Incorporate Partner metadata ingest in ECHO master schedule
  - Generate weekly ingest and holdings reports
- **Support for Metadata Reconciliation**
  - Assist Partner in understanding how to use the API to generate reconciliation data
  - Provide additional reconciliation data as needed (e.g. database "dump" files)

Partner Relations Management

ECHO Ops aims to achieve Partner satisfaction through

- **High Quality User Services**
  - Carefully tracking problems and issues reported by Partners
  - Monitoring time to resolve Partner issues
- **High Quality User Support Products**
  - Working with Partners to identify new/evolving requirements for ECHO user support
  - Conducting surveys of events, materials, and tools
- **System Availability and Performance Monitoring**
  - With feedback to ECHO Dev for HW/SW requirements planning and system evolution
- **Outreach Activities**
  - Promoting project visibility in NASA ESE and external communities
  - Promoting and successfully engaging new Client Partners
  - Advertising Data Partner holdings and availability of new datasets

Data Partner Resources – URLs

- **Project web site** http://www.echo.eos.nasa.gov
- **API documentation** http://api.echo.eos.nasa.gov/echo/message_detail.html
- **Operational system access points**
  - XML Message Test Facility: http://api.echo.eos.nasa.gov/echo/rmi/EchoTestFacility.jsp
  - Client access via SOAP (incl. PUMP): api.echo.eos.nasa.gov/soap/servlet/rpcrouter
  - Metadata FTP: ingest.echo.eos.nasa.gov
  - Browse imagery: browse.echo.eos.nasa.gov
- **Test system access points**
  - XML Message Test Facility: http://beamish.gsfc.nasa.gov:4800/echo/rmi/EchoTestFacility.jsp
  - Client access via SOAP (incl. PUMP): beamish.gsfc.nasa.gov:4800/soap/servlet/rpcrouter

Data Partner Resources – Documents and Tools

- **ECHO Documentation is maintained online at**
  http://www.echo.eos.nasa.gov/echo-docs.shtml
  - ECHO 5.0 Features/Functionalities *(coming soon ECHO 5.0.1)*
  - ECHO API Changes between Version 4.5 and 5.0
  - ECHO 5.0 User's Guide
  - ECHO DTD Tag Directory
  - ECHO Acronym List
- **The Provider User Management Program (PUMP) is available for download from:**
  http://www.echo.eos.nasa.gov/pump_releases/v_1_0_0/install.htm

ECHO Ops – Current Priorities (11/2003)

- **Complete deployment of *ECHO Scout***
  - Monitors overall ECHO system availability to end users in real-time on both the operational and test systems;
  - Alert function notifies ECHO Ops and logs errors with preliminary diagnostic information.
- **Distribute weekly status reports to Data Partners**
- **Finalize Operations Agreements with existing Data Partners**
- **Publish master ingest schedule for Nov-Dec 2003**
- **Assist in ECHO Version 5.0.1 transition to operations**

ECHO Ops - Contact Info

- **Need help from ECHO Ops? Contact us by email at echo-ops@killians.gsfc.nasa.gov -- we'll make sure you get a human to assist you ASAP!**
- **Key ECHO Ops members include**

Medora Macie	GSFC	NASA Project Manager	301-614-6812
Beth Weinstein	GSFC	NASA Project Manager	301-614-5318
Jackie Kendall	SSAI	Task Lead / Coordinator	301-867-2026
Chao-Hsi Chang	GST	Partner Coordinator	301-867-2059
Alex Lai	GST	API Support	301-867-2061
Frank Corprew	GST	Ingest Manager	301-867-2058
Claudia Castaneda	GST	Database Administrator	301-867-2124